

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		Length	DB	ID	Description
		Match	Score				
1	139	100.0	275	5	Q9NIM2	Q9nim2 babesia mic	
2	139	100.0	396	5	Q9NIM4	Q9nim4 babesia mic	
3	60.5	43.5	879	5	Q20227	Q20227 caenorhabdi	
4	54	38.8	1048	3	Q12037	Q12037 saccharomyc	
5	52.5	37.8	661	12	Q9JG50	Q9jg50 ttv-like mic	
6	51	36.7	651	16	Q9ZTW9	Q9ztw9 chlamydia p	
7	50	36.0	196	5	O01439	O01439 caenorhabdi	
8	50	36.0	465	2	Q56267	Q56267 thiobacillu	
9	50	36.0	579	16	Q9Z8K7	Q9z8k7 chlamydia p	
10	50	36.0	923	16	Q9PM90	Q9pm90 campylobact	
11	50	36.0	3268	3	Q93280	Q93280 saccharomyc	
12	49	35.3	273	4	Q9Z951	Q9z951 homo sapien	
13	49	35.3	400	5	Q9VBX9	Q9vbx9 drosophila	
14	49	35.3	803	3	O13657	O13657 schizosacch	
15	48.5	34.9	529	10	Q9FM33	Q9fm33 arabidopsis	
16	48.5	34.9	4833	11	O90VX6	O9gvx6 mus musculu	

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SEROREACTIVE ANTIGEN BN1-17 (FRAGMENT).
OS Babesia microti.
OC Eukaryota; Alveolata; Apicomplexa; Piropalmsida; Babesiidae; Babesia.
OX NCBI_TaxID=5868;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MN1.
RX MEDLINE=20231818; PubMed=10768973;
RA Lodes M.J., Houghton R.L., Bruinsma E.S., Mohamath R., Reynolds L.D.,
RA Benson D.R., Krause P.J., Reed S.G., Persing D.H.;
RT "Serological expression cloning of novel immunoreactive antigens of
RT Babesia microti."
RL Infect. Immun. 68:2783-2790(2000).
DR EMBL: AF206526; AAF68253.1; -
FT NON_TER 396
SQ SEQUENCE 396 AA; 44720 MW; 706E153BA5EB6B7C CRC64;

Query Match 100.0%; Score 139; DB 5; Length 396;
Best Local Similarity 100.0%; Pred. No. 6.8e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKPNTNKSEAKERKSHDTQTQICE 26
Db 133 GKPNTNKSEAKERKSHDTQTQICE 158

RESULT 3
Q20227 ID Q20227 PRELIMINARY; PRT; 879 AA.
AC Q20227;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F40F8.5 PROTEIN.
GN F40F8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC MacDougall R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: 269302; CAA93261.1; -
DR InterPro: IPR003439; ABC_transporter.
DR PROSITE: PS02111; ABC_TRANSPORTER; UNKNOWN.1.
SQ SEQUENCE 879 AA; 99694 MW; BE4403D836B7E3C2 CRC64;

Query Match 43.5%; Score 60.5; DB 5; Length 879;
Best Local Similarity 52.2%; Pred. No. 2.4;
Matches 12; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 5 TNKSEK-AERKSHDTQTQICE 26
Db 3 TNKREKQLEKRECKTQSAVCE 25

RESULT 4
Q12037 ID Q12037 PRELIMINARY; PRT; 1048 AA.
AC Q12037;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN LI309.
GN SDC25 OR SCD25 OR LI309/SCD25 OR YLI016W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALPHA S288C;
RA Purnelle B., Goifeau A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Goifeau A., Purnelle B.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Miosga T., Zimmermann F.K.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=FY23 /RD005 (DERIVED FROM S288C);
RX MEDLINE=96405918; PubMed=8810043;
RA Miosga T., Zimmermann F.K.;
RT "Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on
RT a 43.7 kb fragment of chromosome XII including an open reading frame
RT homologous to the human cystic fibrosis transmembrane conductance
RT regulator protein CFTR."
RL Yeast 12:693-708(1996).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=FY23 /RD005 (DERIVED FROM S288C);
RX MEDLINE=91094833; PubMed=1986220;
RA Danak F., Boy-Marcotte E., Le-Roscouet D., Guilbaud R., Jacquet M.;
RT "SPC25, a CDC25-like gene which contains a RAS-activating domain and
RT is a dispensable gene of Saccharomyces cerevisiae."
RL Mol. Cell. Biol. 11:202-212(1991).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=FY23 /RD005 (DERIVED FROM S288C);
RX MEDLINE=93087480; PubMed=1454790;
RA Rigmaster T.L., Vaughn G.P., Woolford J.L. Jr.;
RT "A putative ATP-dependent RNA helicase involved in Saccharomyces
RT cerevisiae ribosome assembly."
RL Proc. Natl. Acad. Sci. U.S.A. 89:11131-11135(1992).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=FY23 /RD005 (DERIVED FROM S288C);
RX MEDLINE=94375516; PubMed=8089172;
RA Burgess S.M., Delannoy M., Jensen R.E.;
RT "MM1 encodes a mitochondrial outer membrane protein essential for
RT establishing and maintaining the structure of yeast mitochondria."
RL J. Cell Biol. 126:1375-1391(1994).
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=FY23 /RD005 (DERIVED FROM S288C);
RX MEDLINE=95348179; PubMed=7622557;
RA Gammie A.E., Kurihara L.J., Vallee R.B., Rose M.D.;
RT "DNM1, a dynamin-related gene, participates in endosomal trafficking
RT in yeast."
RL J. Cell Biol. 130:553-566(1995).
DR EMBL: X97560; CAA66161.1; -
DR EMBL: Z73121; CAA97461.1; -
DR EMBL: X91488; CAA62775.1; -
DR SGD: S0003939; SDC25.
DR InterPro: IPR000651; RasGEFN.
DR InterPro: IPR001895; RasGEF; 1.
DR Pfam: PF00617; RasGEF; 1.

```

```
DR Pfam: PF00618; RasGEFN: 1.
DR SMART: SM00147; RasGEF: 1.
DR SMART: SM00229; RasGEFN: 1.
DR PROSITE: PS00720; GDS_CD225: 1.
SQ SEQUENCE 1048 AA; 121960 MW; 40A1F12F9B86F458 CRC64;

Query Match 38.8%; Score 54; DB 3; Length 1048;
Best Local Similarity 45.9%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 3 PNTNKSSEKAEKSHDTQTQ 22
Db 1002 PNSNSNKSQEKSRDDQTD 1021

RESULT 5
Q9JG50 SEQUENCE FROM N.A.
ID Q9JG50 PRELIMINARY; PRT; 661 AA.
AC Q9JG50;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ORF1.
OS TTV-like mini virus.
OC Viruses; ssDNA viruses; Circoviridae.
OX NCBI_TaxID=93678;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TLMV-NLC026;
RA Mishihiro S., Hijikata M., Takahashi K.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-TLMV-NLC026;
RX MEDLINE=20428649; PubMed=10971131;
RA Takahashi K., Hijikata M., Samokhvalov E.I., Mishihiro S.;
RT "Full or near full length nucleotide sequences of TT virus variants
(types SANBAN and YONBAN) and the TT virus-like mini virus.";
RL Intervirology 43:119-123(2000).
DR EMBL: AB038630; BAA93609.1; -
DR InterPro: IPR004219; TT_ORF1.
DR Pfam: PF02936; TT_ORF1; 1.
SQ SEQUENCE 661 AA; 78995 MW; 83CE5F93A1AD286A CRC64;

Query Match 37.8%; Score 52.5; DB 12; Length 661;
Best Local Similarity 51.7%; Pred. No. 25;
Matches 15; Conservative 3; Mismatches 6; Indels 5; Gaps 2;

Qy 3 PNTNKSSEKAEKSHDTQTQ-ICE 26
Db 566 PNTNKSSELYCFDERRRQLTKTAQKRICE 594

RESULT 6
Q9Z7W9 SEQUENCE FROM N.A.
ID Q9Z7W9 PRELIMINARY; PRT; 651 AA.
AC Q9Z7W9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE SIMILARITY TO CHLPS INCA (CPJ0585 PROTEIN).
GN CPN0585 OR CPJ0585 OR CP0163.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RN Nat. Genet. 21:385-389(1999).
```

```
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Hincham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae Ar39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL: AE001643; AAD18724.1; -
DR EMBL: AE002177; AAF38042.1; -
DR EMBL: AP002547; BAA98792.1; -
DR TIGR: CP0163; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 75668 MW; 7EAC83171535CAB8 CRC64;

Query Match 36.7%; Score 51; DB 16; Length 651;
Best Local Similarity 38.1%; Pred. No. 40;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 6 NKSEKAEKSHDTQTQICE 26
Db 270 NQSQDIQRAHREASQACE 290

RESULT 7
O01439 SEQUENCE FROM N.A.
ID O01439 PRELIMINARY; PRT; 196 AA.
AC O01439;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 21.5 KDA PROTEIN.
GN B0261.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RA Scheet P., Gattung S.;
RT "The sequence of C. elegans cosmid B0261.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U97016; AAB52352.1; -
KW Hypothetical protein.
SQ SEQUENCE 196 AA; 21516 MW; AF8FFBE84C60EF48 CRC64;
```

Query Match 36.0%; Score 50; DB 5; Length 196;
Best Local Similarity 43.5%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 KPTNTKSEAKRSHDTQTQIBI 24
DB 105 KASDKRKKANKSPRTOEVB 127

RESULT 8

Q56267 ID Q56267 PRELIMINARY; PRT; 465 AA.
AC Q56267;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GLUTAMATE SYNTHASE SMALL SUBUNIT GLTD.
GN GLTD.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33020;
RA Deane S.M., Rawlings D.E.;
RT "The gene for the Thiobacillus ferrooxidans ATCC33020 glutamate
RT synthase (GOGAT) small subunit complements an Escherichia coli gltd
RL mutant";
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -1- COPACTOR: FAD (BY SIMILARITY).
DR EMBL; U36427; AAA79783.1; -;
DR InterPro; IPR000759; Adrnx_reductase.
DR InterPro; IPR001327; FAD_pyr_redux.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00070; pyr_redux; 1.
DR PRINTS; PR00419; ADAROTASE.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 465 AA; 50793 MW; B38EAD4B4D0BEE6A CRC64;

Query Match 36.0%; Score 50; DB 2; Length 465;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 6 NKSEAKRSHDTQTQIBIC 25
DB 73 NRLEEAATLSHTNTLPEIC 92

RESULT 9

Q928K7 ID Q928K7 PRELIMINARY; PRT; 579 AA.
AC Q928K7;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE C082 HYPOTHETICAL PROTEIN.
GN CPN0331 OR CPJ0331 OR CP0426.
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AE001617; AADI8480.1; -;
DR EMBL; AE002204; AAF38269.1; -;
DR EMBL; AP002546; BAA98541.1; -;
DR TIGR; CP0426; -;
KW Complete proteome.
SQ SEQUENCE 579 AA; 61056 MW; A74410B399B9AC7C CRC64;

Query Match 36.0%; Score 50; DB 16; Length 579;
Best Local Similarity 30.8%; Pred. No. 50;
Matches 8; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 GKPNTKSEAKRSHDTQTQICE 26
DB 287 GIPNSNTLEAEKAEKQESRQUSE 312

RESULT 10

Q9PW90 ID Q9PW90 PRELIMINARY; PRT; 923 AA.
AC Q9PW90;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE OXIDOREDUCTASE.
GN CJ1585C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; AL139079; CAB73573.1; -;
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR001467; Molybdopterin.
DR InterPro; IPR001575; Oxid_FAD_bind.
DR InterPro; IPR003170; MurB.
DR Pfam; PF01565; FAD_binding_4; 1.
DR Pfam; PF00037; fer4; 1.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.
DR PROSITE; PS00553; MOLYBDOPTERIN_PROK_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 923 AA; 103909 MW; 06FC04D63259322A CRC64;

Query Match 36.0%; Score 50; DB 16; Length 923;
Best Local Similarity 47.6%; Pred. No. 80;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy	6	NKSEKAERKSHDTQTTEICE	26
		:	
Db	707	NKSEKPNKMHDKRSLOEVFE	727

RESULT 11

Q03280	PRELIMINARY;	PRT: 3268 AA.
ID	Q03280	
AC	Q03280;	
DT	01-NOV-1996 (TrEMBLrel. 01, Created)	
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)	
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)	
DE	D8035.1P (UBIQUITIN LIGASE).	
GN	TOM1 OR D8035.1 OR YDR457W.	
OS	Saccharomyces cerevisiae (Baker's yeast).	
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharo	
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.	
OX	NCBI_TaxID:4932;	
RI	[1]	
RP	SEQUENCE FROM N.A.	
RA	Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles	
RA	Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M,	
RA	Hunnicke-Smith S., Hyman F., Komp C., Lashkari D., Lew H.,	
RA	Mosedale D., Nakahara K., Namath M., Oefner P., Oh C., P	
RA	Roberts D., Schramm S., Schroeder M., Shogren T., Shroff	
RA	Winant A., Yelton M., Botstein D., Davis R.W.;	
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.	

RP	SEQUENCE FROM N.A.
RA	Dietrich F.S.;
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RP	{3}
RP	SEQUENCE FROM N.A.
RA	Jia Y., Cherry J.M.;
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RP	{4}
RP	SEQUENCE OF 2156-3268 FROM N.A.
RA	Utsugi T., Kikuchi Y.;
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR	EMBL; U30050; AAB64910.1; -.
DR	EMBL; D63905; BAA21482.1; -.
DR	SGD; S0002865; TOM1.
DR	InterPro; IPR000569; HECT.
DR	Pfam; PF00632; HECT; 1.
DR	SMART; SM00119; HECTc; 1.
DR	PROSITE; PS50237; HECT; 1.
KW	Ligase.

Query Match 36.0%; Score 50; DB 3; Length 3268;
Best Local Similarity 44.0%; Pred. No. 2.9e+02;
Matches 11: Conservative 6; Mismatches 6; Indels

QY 2 KPNTNKSEKAERKSHDTQTTEICE 26
 I I I :: : I I I I I I I I
 Db 226 KNNINETOTSIKKV--TMTTOELCE 248

RESULT 12

Q92951	PRELIMINARY;	PRT;	273 AA.
ID	Q92951		
AC	Q92951;		
DT	01-FEB-1997 (TREMBLrel. 02, Created)		
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 02, Last annotation update)		
DE	ZINC FINGER PROTEIN.		
OS	Homo sapiens (Human).		
OC	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_Taxid=9606;		
	[1]		
RN	SEQUENCE FROM N.A.		

RX MEDLINE=98135459; PubMed=9475118;
 RA Drew P.D., Gado A.M., Canning R.D., Nagle J.W., Dehejia A.M.,
 RA Polymopoulos M.H., Bidison W.E., Jacobson S., Becker K.G.;
 RT "C2H2-546: a zinc finger protein differentially expressed in HTLV-1
 RT infected T cells";
 RL J. Neurovirol. 3:455-459(1997).
 CC J-1. SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY)

Query Match 35.3%; Score 49; DB 4; Length 273;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 10; Conservative 5; Mismatches 5; Indels

Qy 4 NTNKSEKAERKSHDTQTQE 23
| : | | | : | : | :
Db 48 NSFREKLEOKSPDSTLOE 67

RESULT 13

Q9VB99	PRELIMINARY; PRFT; 400 AA.
ID	Q9VB99
AC	Q9VB99;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	CGI3649 PROTEIN.
GN	CGI3649.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE=20136006; PubMed=10731132;
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.V., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke A., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Hougk C.,
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwac J.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reipert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003750; AAF56387.1; -.
DR FlyBase; FBgn0035267; CGI3649.
DR InterPro; IPR003889; FYRICH_C.
DR SMART; SM003888; FYRICH_N.
DR SMART; SM00542; FYRC; 1.
DR SMART; SM00541; FYRN; 1.
SQ SEQUENCE 400 AA; 45117 MW; 71CA68B1D6B50E3F CRC64;

Query Match 35.3%; Score 49; DB 5; Length 400;
Best Local Similarity 42.9%; Pred. NO. 47;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PNTNKSKEAERKSHDTQTQEE 23
|||:::|||||
DB 320 PNRNEEQATRRSESTQAE 340

RESULT 14
Q13657
ID O13657 PRELIMINARY; PRT; 803 AA.
AC O13657;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PROBABLE MEMBRANE PROTEIN YOL130W.
GN P1066 OR SPBC27B12.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972 H-;
RA Kushida N., Yamazaki S., Tanaka T., Jinno K., Haikawa Y., Yamazaki J.,
RA Yamamoto S., Sekine M., Oguchi A., Nagai Y., Sakai M., Aoki K.,
RA Ogura K., Otsuka R., Kudoh Y., Yanagida M., Machida M., Zhang M.Q.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,
RA Dusterhoeft A.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB004538; BAA21447.1; -.
DR EMBL; AL021766; CAAL6907.1; -.
DR EMBL; AB004539; BAA21448.1; -.
DR InterPro; IPR002523; CoRa.
DR Pfam; PF01544; CoRa; 1.
KW Hypothetical protein.
SQ SEQUENCE 803 AA; 89396 MW; 5E1DB270D9D83572 CRC64;

Query Match 35.3%; Score 49; DB 3; Length 803;
Best Local Similarity 34.8%; Pred. NO. 96;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 KPNTNKSEKAERKSHDTQTQEI 24
|||:::| | |::| |
DB 150 KPQSNKHKRRVRKHPKSTLEV 172

RESULT 15
Q9FNJ3
```

```
ID Q9FNJ3 PRELIMINARY; PRT; 529 AA.
AC Q9FNJ3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K19P17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned pl clones.";
RL DNA Res. 4:401-414(1997).
DR EMBL; AB007644; BAB10723.1; -.
SQ SEQUENCE 529 AA; 59398 MW; 6A85A27764416EEB CRC64;

Query Match 34.9%; Score 48.5; DB 10; Length 529;
Best Local Similarity 38.5%; Pred. NO. 74;
Matches 10; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 1 GKPTNKSEKAERKSHDTQTQICE 26
|:::|||||::|
DB 312 GEPMQIESEKVESQDHDLN-EOKLCD 336

Search completed: October 11, 2002, 20:40:04
Job time : 89 secs
```


553079

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 19:21:53 ; Search time 33 Seconds
(without alignments)
30.506 Million cell updates/sec

Title: US-09-794-764-195
Perfect score: 139
Sequence: 1 GKPTNKSEAKERSHDTQTQBICE 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	56	40.3	511	1 MDJ1_YEAST	P35191 saccharomyc
2	54	38.8	1158	1 RL14_HUMAN	P48552 homo sapien
3	54	38.8	1253	1 SC25_YEAST	P14771 saccharomyc
4	50	36.0	1409	1 AEX3_CAEEL	O02626 caenorhabdi
5	49	35.3	210	1 OM22_ARATH	P82873 arabidopsis
6	47	33.8	120	1 TRI7_HUMAN	Q15651 homo sapien
7	47	33.8	178	1 EYX5_HUMAN	Q96db9 homo sapien
8	47	33.8	369	1 VP6_AHSV3	Q84909 african hor
9	47	33.8	369	1 VP6_AHSV6	Q84913 african hor
10	47	33.8	2363	1 SPCO_MOUSE	Q62261 mus musculu
11	46	33.1	533	1 YN34_YEAST	P48565 saccharomyc
12	45	32.4	148	1 CALM_EUGGR	P11118 euglena gra
13	45	32.4	263	1 SGS3_DROYA	P13728 drosophila
14	45	32.4	798	1 YC00_MYCPN	Q50288 mycoplasma
15	45	32.4	843	1 YMS1_YEAST	Q05050 saccharomyc
16	45	32.4	929	1 RBMA_HUMAN	P98175 homo sapien
17	45	32.4	1020	1 NFH_HUMAN	P12036 homo sapien
18	44.5	32.0	906	1 CENC_MOUSE	P49452 mus musculu
19	44	31.7	60	1 RL32_DEIRA	P49228 deinococcus
20	44	31.7	327	1 HUNB_MANSE	Q35514 manduca sex
21	44	31.7	328	1 VP6_BTVIS	Q03328 bluetongue
22	44	31.7	343	1 SLAM_MOUSE	Q9qu44 mus musculu
23	44	31.7	1803	1 YJL3_YEAST	P47024 saccharomyc
24	43.5	31.3	178	1 YK01_ARCFU	Q28278 archaeoglob
25	43	30.9	52	1 RK32_CYACA	O19894 cyanidium c
26	43	30.9	56	1 RL32_MYCPN	P75238 mycoplasma
27	43	30.9	63	1 RL32_XYLFA	Q9pcg5 xyliella fas
28	43	30.9	325	1 LXBI_PHOLE	P09141 photobacter
29	43	30.9	406	1 SNX6_HUMAN	Q9unh7 homo sapien
30	43	30.9	406	1 YDLB_SCHPO	P87126 schizosacch
31	43	30.9	436	1 Y868_CHLMU	Q9pjgl chlamydia m
32	43	30.9	547	1 OM6C_CHLTR	P26758 chlamydia t
33	43	30.9	547	1 OM6D_CHLTR	P18151 chlamydia t

RESULT 1									
ID	MDJ1_YEAST								
AC	P35191;								
DT	01-FEB-1994 (Rel. 28, Created)								
DT	01-FEB-1994 (Rel. 28, Last sequence update)								
DT	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	MDJ1 protein, mitochondrial precursor.								
GN	MDJ1 OR YFL016C.								
OS	Saccharomyces cerevisiae (Baker's yeast).								
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;								
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.								
OX	NCBI_TaxID=4932;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=S288C / AB972;								
RA	Rowley N.K., Prip-Buus C., Westermann B., Brown C.M., Schwarz E.,								
RA	Barrell B.G., Neupert W.;								
RL	Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=S288C / AB972;								
RA	Barrell B.G., Churcher C., Rajandream M.A.;								
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=S288C / AB972;								
RC	MEDLINE=95400292; PubMed=7670463;								
RA	Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,								
RA	Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,								
RA	Yamazaki M., Tashiro H., Eki T.;								
RT	"Analysis of the nucleotide sequence of chromosome VI from								
RL	Saccharomyces cerevisiae.";								
CC	Nat. Genet. 10:261-268(1995).								
CC	-1- FUNCTION: PLAYS A ROLE IN MITOCHONDRIAL BIOGENESIS AND PROTEIN								
CC	FOLDING.								
CC	-1- SUBCELLULAR LOCATION: Mitochondrial.								
CC	-1- SIMILARITY: CONTAINS 1 J DOMAIN.								
CC	-1- SIMILARITY: CONTAINS 1 CR DOMAIN.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC	-----								
DR	EMBL; Z28336; CAA82189.1; -;								
DR	EMBL; Z46255; CAA86351.1; -;								
DR	EMBL; D50617; BAA09222.1; -;								
DR	EMBL; D44596; BAA08001.1; -;								
DR	PIR; S38898; S38898.								
DR	PIR; S48320; S48320.								
DR	HSSP; P08622; IBOZ.								
DR	SGD; S0001878; MDJ1.								

P23603 chlamydia t
P21354 chlamydia t
Q61909 mus musculu
Q06455 homo sapien
Q58192 methanococc
O15234 homo sapien
Q39v04 mus musculu
P70501 rattus norv
Q00132 ictaluriid h
Q99mw1 mus musculu
Q9pmq3 campylobact
P10481 clostridium

ALIGNMENTS

[illegible]

```

RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Eguu P., Feidblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.-X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
CC -!- FUNCTION: CENTRAL COMPONENT OF THE RECEPTOR COMPLEX RESPONSIBLE
CC FOR THE RECOGNITION AND TRANSLOCATION OF CYTOSOLICALLY SYNTHESIZED
CC MITOCHONDRIAL PREPROTEINS. TOGETHER WITH TOM22 FUNCTIONS AS THE
CC TRANSIT PEPTIDE RECEPTOR AT THE SURFACE OF THE MITOCHONDRION
CC OUTER MEMBRANE AND FACILITATES THE MOVEMENT OF PREPROTEINS INTO
CC THE TRANSLOCATION PORE.
CC -!- SUBUNIT: FORMS PART OF MITOCHONDRIAL RECEPTOR COMPLEX.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- MISCELLANEOUS: THERE ARE FOUR GENES (TOM20-1, TOM20-2, TOM20-3 AND
CC TOM20-4) WHICH ENCODE MITOCHONDRIAL IMPORT RECEPTOR SUBUNITS
CC TOM20.
CC -!- SIMILARITY: BELONGS TO THE TOM20 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ296024; CAC14429.1; -.
DR EMBL: AC004557; AAF99745.1; -.
KW Transport; Protein transport; Outer membrane; Mitochondrion;
KW Transmembrane.
FT DOMAIN 1 178 INTERMEMBRANE (POTENTIAL).
FT TRANSMEM 179 199 POTENTIAL.
FT DOMAIN 200 210 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 11 11 F -> L (IN REF. 2).
FT CONFLICT 168 168 MISSING (IN REF. 2).
SQ SEQUENCE 210 AA; 23204 MW; 65EA3327249A62E2 CRC64;

Query Match 35.3%; Score 49; DB 1; Length 210;
Best Local Similarity 28.0%; Pred. No. 7.3;
Matches 7; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKAEKSHDQTQTQEI 25
DB 160 GMASNVQSQSKKKRNTFTYDVC 184

RESULT 6
TR17_HUMAN
ID TR17_HUMAN STANDARD; PRT; 120 AA.
AC Q15651;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thyroid receptor interacting protein 7 (TRIP7) (Fragment).
GN TRIP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=95295737; PubMed=7776974;
RA Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RT "Two classes of proteins dependent on either the presence or absence
RT of thyroid hormone for interaction with the thyroid hormone
RT receptor.";
RL Mol. Endocrinol. 9:243-254(1995).
CC -!- FUNCTION: INTERACTS, IN VITRO, WITH THE LIGAND BINDING DOMAIN OF
CC THE THYROID RECEPTOR (TR). TRIP7 REQUIRES THE PRESENCE OF THYROID
CC HORMONE FOR ITS INTERACTION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN KIDNEY, SKELETAL
CC MUSCLE AND HEART. LOWER LEVELS FOUND IN LUNG, LIVER AND PANCREAS.
CC -!- SIMILARITY: BELONGS TO THE HMG-14/HMG-17 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L40357; AAA73877.1; -.
DR MIM: 604502; -.
DR InterPro: IPR000079; Highmoblty_14_17.
DR Pfam: PF01101; HMG14_17; 1.
DR SMART: SM00527; HMG17; 1.
DR PROSITE: PS00355; HMG14_17; 1.
KW Nuclear protein; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 120 AA; 13016 MW; 2B7A0B9D352BF78F CRC64;

Query Match 33.8%; Score 47; DB 1; Length 120;
Best Local Similarity 37.5%; Pred. No. 7.7;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKAEKSHDQTQTQEI 24
DB 92 GREGTAPSENGETKAEQAQKTESV 115

RESULT 7
FXYS_HUMAN
ID FXYS_HUMAN STANDARD; PRT; 178 AA.
AC Q96DB9; Q9P039; Q9HC34;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE FXYS domain-containing ion transport regulator 5 precursor (HSPC113).
GN FXYS OR IWU1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).

```


Spectrin beta chain, brain 1 (Spectrin, non-erythroid beta chain 1)
 (Beta-II spectrin) (Fodrin beta chain).
 GN SPTN1 OR SPTB2 OR SPNB2 OR SPNB-2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C;
 RX MEDLINE=93240985; PubMed=8479293;
 RA Ma Y., Zimmer W.E., Riederer B.M., Goodman S.R.;
 RT "The complete amino acid sequence for brain beta spectrin (beta
 fodrin): relationship to globin sequences.";
 RL Brain Res. Mol. Brain Res. 18:87-89(1993).
 RN [2]
 RP STRUCTURE BY NMR OF 2199-2304.
 RX MEDLINE=94268558; PubMed=8208297;
 RA Macias M.J., Musacchio A., Ponstingl H., Nilges M., Saraste M.,
 RA Oschkinat H.;
 RT "Structure of the pleckstrin homology domain from beta-spectrin.";
 RL Nature 369:675-677(1994).
 RN [3]
 RP STRUCTURE BY NMR OF 2199-2304.
 RX MEDLINE=97342712; PubMed=9199409;
 RA Nilges M., Macias M.J., O'Donoghue S.I., Oschkinat H.;
 RT "Automated NOESY interpretation with ambiguous distance restraints:
 the refined NMR solution structure of the pleckstrin homology domain
 from beta-spectrin.";
 RT J. Mol. Biol. 269:408-422(1997).
 RL [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 2199-2304.
 RX MEDLINE=96030773; PubMed=7588597;
 RA Hyvoenen M., Macias M.J., Nilges M., Oschkinat H., Saraste M.,
 RA Wilmanns M.;
 RT "Structure of the binding site for inositol phosphates in a PH
 domain.";
 RL EMBO J. 14:4676-4681(1995).
 CC -!- FUNCTION: FODRIN, WHICH SEEMS TO BE INVOLVED IN SECRETION,
 CC INTERACTS WITH CALMODULIN IN A CALCIUM-DEPENDENT MANNER AND IS
 CC THIS CANDIDATE FOR THE CALCIUM-DEPENDENT MOVEMENT OF THE
 CC CYTOSKELETON AT THE MEMBRANE.
 CC -!- SUBUNIT: LIKE ERYTHROCYTE SPECTRIN, THE SPECTRIN-LIKE PROTEINS
 CC ARE CAPABLE TO FORM DIMERS WHICH CAN FURTHER ASSOCIATE TO
 CC TETRAMERS.
 CC -!- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -!- SIMILARITY: CONTAINS 17 SPECTRIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M74773; AAC42040.1; -
 CC PDB; 1BTN; 08-MAR-96.
 CC PDB; 1MPH; 16-JUN-97.
 CC MGI; 98388; Spnb2.
 CC InterPro; IPR001589; Actinin_act_bind.
 CC InterPro; IPR001715; Calponin_hom.
 CC InterPro; IPR001849; PH.
 CC InterPro; IPR002017; Spectrin.
 CC InterPro; IPR001605; Spectrin_PH.
 CC Pfam; PF00307; CH; 2.
 CC Pfam; PF00169; PH; 1.
 CC Pfam; PF00435; spectrin; 18.
 CC PRINTS; PR00683; SPECTRINPH.
 CC SMART; SM00033; CH; 2.
 CC SMART; SM00233; PH; 1.

DR SMART; SM00150; SPEC; 16.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 KW Cytoskeleton; Membrane; Repeat; Actin-binding; Capping protein;
 3D-structure.
 FT DOMAIN 1 275 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 54 158 CH 1.
 FT DOMAIN 173 275 CH 2.
 FT REPEAT 276 384 SPECTRIN 1.
 FT REPEAT 385 498 SPECTRIN 2.
 FT REPEAT 499 608 SPECTRIN 3.
 FT REPEAT 609 714 SPECTRIN 4.
 FT REPEAT 715 819 SPECTRIN 5.
 FT REPEAT 820 925 SPECTRIN 6.
 FT REPEAT 926 1032 SPECTRIN 7.
 FT REPEAT 1033 1139 SPECTRIN 8.
 FT REPEAT 1140 1245 SPECTRIN 9.
 FT REPEAT 1246 1350 SPECTRIN 10.
 FT REPEAT 1351 1462 SPECTRIN 11.
 FT REPEAT 1463 1562 SPECTRIN 12.
 FT REPEAT 1563 1668 SPECTRIN 13.
 FT REPEAT 1669 1775 SPECTRIN 14.
 FT REPEAT 1776 1881 SPECTRIN 15.
 FT REPEAT 1882 1987 SPECTRIN 16.
 FT REPEAT 1988 2132 SPECTRIN 17.
 FT DOMAIN 2196 2306 PH.
 SQ SEQUENCE 2363 AA; 274420 MW; 64C9E4BD26B8C7B8 CRC64;
 Query Match 33.8%; Score 47; DB 1; Length 2363;
 Best Local Similarity 37.5%; Pred. No. 1.7e+02;
 Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 3 PNTKSKERAKSHDTQTQICE 26
 DB 2105 PNTKVSSEAEQQWDTSRGDQVSQ 2128
 RESULT 11
 YN34_YEAST
 ID YN34_YEAST STANDARD; PRT; 533 AA.
 AC P48565;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hypothetical 61.5 kDa protein in CLA4-PUS4 intergenic region.
 GN YNL294C OR N0466.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=96137033; PubMed=8553702;
 RA Maurer K.C.T., Urbanus J.H.M., Planta R.J.;
 RT "Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV
 RT carrying a ribosomal protein gene cluster, the genes encoding a
 RT plasma membrane protein and a subunit of replication factor C, and a
 RT novel putative serine/threonine protein kinase gene.";
 RL Yeast 11:1303-1310(1995).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U23084; AAC49105.1; -

```
DR EMBL; Z71570; CAA96212.1; -.
KW SGD; S0005238; YNL294C.
DR Hypothetical protein: Transmembrane.
FT TRANSMEM 95 115 POTENTIAL.
FT TRANSMEM 128 148 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT TRANSMEM 248 268 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
SQ SEQUENCE 533 AA; 61532 MW; 9492A18512F399CC CRC64;

Query Match 33.1%; Score 46; DB 1; Length 533;
Best Local Similarity 35.0%; Pred. No. 49;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 7 KSEAERKSHDTQTQTEICE 26
      :||| :||| :||| :|||
Db 376 ESQASTDRDHTSSNSEVCD 395

RESULT 12
CALM_EUGGR
ID CALM_EUGGR STANDARD; PRT; 148 AA.
AC P11118;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Calmodulin.
OS Euglena gracilis.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE.
RC STRAIN=Z;
RX MEDLINE=92241300; PubMed=1572365;
RA Toda H., Yazawa M., Yagi Y.;
RT "Amino acid sequence of calmodulin from Euglena gracilis.";
RL Eur. J. Biochem. 205:653-660(1992).
CC -!- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
CC ENZYMES BY CA(++) . AMONG THE ENZYMES TO BE STIMULATED BY THE
CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
CC PHOSPHATASES.
CC -!- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
CC SITES.
CC -!- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
DR PIR; JK0008; MCEG.
DR PIR; S21212; S21212.
DR HSSP; P02593; LCDM.
DR InterPro; IPR002048; EF-hand.
DR SMART; SM00054; EFh; 4.
DR PROSITE; PS00018; EF_HAND; 4.
KW Calcium-binding; Repeat; Acetylation; Methylation.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 115 115 METHYLATION.
FT MOD_RES 148 148 METHYLATION.
FT CA_BIND 20 31 EF-HAND 1.
FT CA_BIND 56 67 EF-HAND 2.
FT CA_BIND 93 104 EF-HAND 3.
FT CA_BIND 129 140 EF-HAND 4.
SQ SEQUENCE 148 AA; 16724 MW; 82A1E48108638455 CRC64;

Query Match 32.4%; Score 45; DB 1; Length 148;
Best Local Similarity 64.3%; Pred. No. 18;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 RSHDTPQTQTEICE 26
      ||||| :|||
Db 74 RKMHTDTEERKE 87

RESULT 13
SGS3_DROYA
```

```
ID SGS3_DROYA STANDARD; PRT; 263 AA.
AC P13728;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Salivary glue protein Sgs-3 precursor.
GN SGS3.
OS Drosophila yakuba (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88332966; PubMed=3138416;
RA Martin C.H., Mayeda C.A., Meyerowitz E.M.;
RT "Evolution and expression of the Sgs-3 glue gene of Drosophila.";
RL J. Mol. Biol. 201:273-287(1988).
CC -!- DEVELOPMENTAL STAGE: PRODUCED BY THIRD-INSTAR LARVAE.
DR PIR; S01360; S01360.
DR FlyBase; FBgn0013172; Dyak\SGS3.
KW Repeat; Signal;
FT SIGNAL 1 23
FT CHAIN 24 263
SQ SEQUENCE 263 AA; 28392 MW; C0C5246B482A261C CRC64;

Query Match 32.4%; Score 45; DB 1; Length 263;
Best Local Similarity 40.0%; Pred. No. 32;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 4 NTNKSEAKERKSHDTQTQTE 23
      :||| :||| :||| :|||
Db 168 HTTKSTTSKRPTHETTTSK 187

RESULT 14
YC00_MYCPN
ID YC00_MYCPN STANDARD; PRT; 798 AA.
AC Q50288;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein MPN200 precursor (GT9_orf798).
GN MPN200 OR MP631.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=96177562; PubMed=8604303;
RA Hilbert H., Himmelfreich R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RL cluster of ribosomal protein genes.";
RN Nucleic Acids Res. 24:628-639(1996).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -!- SIMILARITY: BELONGS TO THE MG185 / MG260 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U34795; AAC43681.1; -
DR EMBL; AE000060; AAB96279.1; -
KW Hypothetical protein; Lipoprotein; Membrane; Signal;
KW Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 798 HYPOTHETICAL LIPOPROTEIN MPN200.
FT LIPID 23 23 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 798 AA; 87628 MW; 3E2471D7EF0A3CA0 CRC64;

Query Match 32.4%; Score 45; DB 1; Length 798;

Best Local Similarity 40.9%; Pred. No. 1e+02; Mismatches 8; Indels 0; Gaps 0;

Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 KPTNKKSEAKRSHDTQTQEE 23

DB 229 KKEKKSTRKRATEGTQTKE 250

RESULT 15

YMSL_YEAST

ID YMSL_YEAST STANDARD; PRT; 843 AA.
AC Q05050;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 93.3 kba protein in TAP42-CYK2 intergenic region.
GN YMR031C OR YM9973.04C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YKL050C.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z49213; CAA89146.1; -
DR SGD; S0004633; YMR031C.

KW Hypothetical protein.

SQ SEQUENCE 843 AA; 93345 MW; 0068D65A229B941D CRC64;

Query Match

Best Local Similarity 32.4%; Score 45; DB 1; Length 843;

Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 8 SEKAERKSHDTQTQEE 23

DB 88 SDKAANLAHDNQTVE 103

Search completed: October 11, 2002, 20:38:24

Job time : 35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 20:28:08 ; Search time 48 Seconds

(without alignments)
52.048 Million cell updates/sec

Title: US-09-794-764-195

Perfect score: 139

Sequence: 1 GKPTNKSEAKERKSHDTQTQICE 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.5	43.5	879	2 T22033	hypothetical prote
2	56	40.3	511	2 S38898	heat shock protein
3	54	38.8	1048	2 S64758	SCD25 protein (ver
4	54	38.8	1250	2 S14177	SCD25 protein (ver
5	51	36.7	651	2 F85663	hypothetical prote
6	51	36.7	651	2 A72060	hypothetical prote
7	50	36.0	196	2 T25445	hypothetical prote
8	50	36.0	465	2 JC5184	glutamate synthase
9	50	36.0	471	2 AD2545	hypothetical prote
10	50	36.0	579	2 D72092	conserved hypotet
11	50	36.0	579	2 C86532	C7082 hypothetical
12	50	36.0	923	2 G81253	probable oxidoredu
13	50	36.0	1409	2 T37188	presynaptic activi
14	50	36.0	3268	2 S69625	hypothetical prote
15	49	35.3	803	2 T40036	hypothetical prote
16	47.5	34.2	211	1 E69844	GTP pyrophosphokin
17	47	33.8	453	2 F71374	probable hemolysin
18	47	33.8	537	2 A86444	probable RNA helic
19	47	33.8	635	2 F71621	hypothetical prote
20	47	33.8	950	2 T09076	hypothetical prote
21	47	33.8	1097	2 T45622	hypothetical prote
22	46.5	33.5	65	2 G81292	hypothetical prote
23	46.5	33.5	1076	2 F96831	hypothetical prote
24	46	33.1	57	2 T10457	lipoprotein L - ps
25	46	33.1	225	2 T28053	hypothetical prote
26	46	33.1	283	2 T23785	hypothetical prote
27	46	33.1	383	2 H90266	conserved hypotet
28	46	33.1	533	2 S63270	probable membrane
29	46	33.1	608	2 T25572	hypothetical prote

30	46	33.1	761	2 B69797	transcription regu
31	46	33.1	788	2 C95046	helicase, probable
32	46	33.1	831	2 T49758	related to cytochr
33	45.5	32.7	631	2 T13115	protein gp29 - pha
34	45.5	32.7	881	2 T33810	hypothetical prote
35	45.5	32.7	890	2 T34243	hypothetical prote
36	45.5	32.7	1230	2 T19899	hypothetical prote
37	45	32.4	148	1 MCEG	calmodulin - Eugle
38	45	32.4	263	2 S01360	salivary glue prot
39	45	32.4	310	2 T47695	hypothetical prote
40	45	32.4	333	2 T19313	hypothetical prote
41	45	32.4	542	2 T46464	hypothetical prote
42	45	32.4	543	2 T02022	hypothetical prote
43	45	32.4	749	2 A45294	Balbani ring 2.1
44	45	32.4	761	2 S66719	hypothetical prote
45	45	32.4	798	2 S62791	probable lipoprote

ALIGNMENTS

RESULT 1

T22033

hypothetical protein F40F8.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22033

R:MacDougall, R.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z19505

A:Accession: T22033

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-879 <WIL>

A:Cross-references: EMBL:Z69302; PIDN:CAA93261.1; GSPDB:GN000020; CESP:F40F8.5

A:Experimental source: clone F40F8

C:Genetics:

A:Gene: CESP:F40F8.5

A:Map position: 2

A:Introns: 63/1; 149/1; 173/1; 217/1; 262/1; 306/1; 347/1; 381/2; 422/2; 440/3; 461/3

Query Match 43.5%; Score 60.5; DB 2; Length 879;

Best Local Similarity 52.2%; Pred. No. 1.4;

Matches 12; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 5 TNKSEK-AERKSHDTQTQICE 26

Db 3 TNKREKQLERECHQKTSQAVCE 25

RESULT 2

S38898

heat shock protein MDJ1 precursor - yeast (Saccharomyces cerevisiae)

N:Alternate names: MDJ1 protein; protein YFL016c

C:Species: Saccharomyces cerevisiae

C>Date: 31-Dec-1993 #sequence_revision 27-Jan-1995 #text_change 20-Jun-2000

C:Accession: S38898; S48320; A53500; S62297

R:Rowley, N.K.; Prip-Buus, C.; Westermann, B.; Brown, C.M.; Schwarz, E.; Barrell, B.G.

submitted to the EMBL Data Library, November 1993

A:Description: Mdj1p, a novel DnaJ homologue of Saccharomyces cerevisiae, plays a rol

A:Reference number: S38898

A:Accession: S38898

A:Molecule type: DNA

A:Residues: 1-511 <ROW>

A:Cross-references: EMBL:Z28336; NID:q431909; PIDN:CAA82189.1; PID:q431910

R:Churcher, C.

submitted to the EMBL Data Library, September 1994

A:Reference number: S48310

A:Accession: S48320

A:Molecule type: DNA

A:Residues: 1-511 <CHU>

A:Cross-references: EMBL:Z46255; NID:q559925; PIDN:CAA86351.1; PID:q559936; MTPS:YFLO

R:Rowley, N.; Prip-Buus, C.; Westermann, B.; Brown, C.; Schwarz, E.; Barrell, B.; Neu


```

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25445
R;Scheet, P.; Gattung, S.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid B0261.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: F86563
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise, N.
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: F86563
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-651 <SPO>
A;Cross-references: GB:BA000008; NID:g8978957; PIDN:BAA98792.1; GSPDB:GN00142
C;Genetics:
A;Gene: CPJ0585

Query Match 36.7%; Score 51; DB 2; Length 651;
Best Local Similarity 38.1%; Pred. No. 24;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 6 NKSEKAERKSHDTQTQICE 26
I:|:| :|:| :|:| :|:|
Db 270 NQSQDKIQRAHDREASQRACE 290

RESULT 6
A72060
hypothetical protein CP0163 [imported] - Chlamydophila pneumoniae (strains CWL029 and AR39)
N;Alternate names: hypothetical protein cpn0585
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: A72060; B81607
R;Kalmán, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606
A;Accession: A72060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-651 <ARN>
A;Cross-references: GB:AE001643; GB:AE001363; NID:g4376876; PIDN:AAD18724.1; PID:g4376876
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.; C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255
A;Accession: B81607
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-651 <REA>
A;Cross-references: GB:AE002177; GB:AE002161; NID:g7189090; PIDN:AAF38042.1; PID:g7189090
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CPn0585; CP0163

Query Match 36.7%; Score 51; DB 2; Length 651;
Best Local Similarity 38.1%; Pred. No. 24;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 6 NKSEKAERKSHDTQTQICE 26
I:|:| :|:| :|:| :|:|
Db 270 NQSQDKIQRAHDREASQRACE 290

RESULT 7
T25445
hypothetical protein B0261.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

```

```

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25445
R;Scheet, P.; Gattung, S.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid B0261.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: T25445
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-196 <SCH>
A;Cross-references: EMBL:U97016; PIDN:AAB52352.1; GSPDB:GN00019; CESP:B0261.5
A;Experimental source: strain Bristol N2; clone B0261
C;Genetics:
A;Gene: CESP:B0261.5
A;Map position: 1
A;Introns: 162/3

Query Match 36.0%; Score 50; DB 2; Length 196;
Best Local Similarity 43.5%; Pred. No. 9.9;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 2 KPNTNKSEKAERKSHDTQTQEI 24
I:|:| :|:| :|:| :|:|
Db 105 KASDKRKKKANKSPRTOETQEV 127

RESULT 8
JC5184
glutamate synthase (GOGAT) (EC 1.4.1.1-) small chain - Thiobacillus ferrooxidans
C;Species: Thiobacillus ferrooxidans
C;Date: 16-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 29-Sep-1999
C;Accession: JC5184
R;Deane, S.M.; Rawlings, D.E.
Gene 177, 261-263, 1996
A;Title: Cloning and sequencing of the gene for the Thiobacillus ferrooxidans ATCC330
A;Reference number: JC5184; MUID:97080532
A;Accession: JC5184
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-465 <DEA>
A;Cross-references: GB:U36427; NID:g1033069; PIDN:AAA79783.1; PID:g1033071
A;Experimental source: strain ATCC33020
C;Genetics:
A;Gene: gltD
C;Superfamily: glutamate synthase small chain
C;Keywords: iron-sulfur protein; metalloprotein; oxidoreductase
E;151-156,291-296/Region: glycine-rich
F;427-437/Region: FAD-binding site; iron-sulfur clusters (Cys) (covalent) #stat
F;45,48,53,57,92,96,102,106/Binding site:

Query Match 36.0%; Score 50; DB 2; Length 465;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 6 NKSEKAERKSHDTQTQICE 25
I:|:| :|:| :|:| :|:|
Db 73 NRLEEAATLSHOTNLTPEIC 92

RESULT 9
AD2545
hypothetical protein alr7647 [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC7
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AD2545
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2545
A;Status: preliminary

```


hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C;Accession: S69625
R;Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A;Description: The sequence of S. cerevisiae cosmids 9410, 8035, 8166, and 9787.
A;Reference number: S69554
A;Accession: S69625
A;Molecule type: DNA
A;Residues: 1-3268 <DIE>
A;Cross-references: EMBL:U33050; NID:g927726; PIDN:AB64910.1; PID:g927738; MIPS:YDR457w
C;Genetics:
A;Gene: SGD:TOM1
A;Cross-references: SGD:S0002855; MIPS:YDR457w
A;Map position: 4R

Query Match 36.0%; Score 50; DB 2; Length 3268;
Best Local Similarity 44.0%; Pred. No. 1.6e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 2 KPNTNKSEKAERKSHDTQTQICE 26
|||:::|:| |||||
Db 226 KNNINETSIIKV--TMTTQELCE 248

RESULT 15
T40036
hypothetical protein SPBC27B12.12c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40036
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z21900
A;Accession: T40036
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-803 <WOO>
A;Cross-references: EMBL:AL021766; PIDN:CAA16907.1; GSPDB:GN00067; SPDB:SPBC27B12.12c
C;Genetics:
A;Gene: SPDB:SPBC27B12.12c
A;Map position: 2

Query Match 35.3%; Score 49; DB 2; Length 803;
Best Local Similarity 34.8%; Pred. No. 56;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 KPNTNKSEKAERKSHDTQTQEI 24
|||:::|:| |||||
Db 150 KPQSNKKRGRRVKHSPKSTLEV 172

Search completed: October 11, 2002, 20:41:05
Job time : 51 secs

THIS PAGE BLANK (USPTO)

8/5/30/91

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 18:23:07 ; Search time 62 Seconds
(without alignments)
46.579 Million cell updates/sec

Title: US-09-794-764-195
Perfect score: 139
Sequence: 1 GKPNTNKAERKSHDTQTQICE 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032302.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	100.0	275	19 AAW56303	Babesia microti BM
2	139	100.0	275	20 AAY24365	Babesia microti an
3	139	100.0	275	21 AAB30217	B. microti clone a
4	139	100.0	445	19 AAW56298	Babesia microti BM
5	139	100.0	445	20 AAY24358	Babesia microti an
6	139	100.0	445	21 AAB30207	B. microti BMNI-17
7	139	100.0	666	21 AAB30230	B. microti MN-10/B
8	139	100.0	1132	21 AAB30231	B. microti MN-10/B
9	63	45.3	32	20 AAY24359	Babesia microti an
10	63	45.3	32	21 AAB30208	B. microti BMNI-17
11	54	38.8	600	22 AAG75092	Human colon cancer

12	54	38.8	1048	21 AAY68823	Amino acid sequenc
13	51	36.7	212	20 AAY35210	Chlamydia pneumoni
14	50	36.0	584	20 AAY34934	Chlamydia pneumoni
15	49.5	35.6	523	21 AAY56856	Morbillivirus anti
16	49.5	35.6	776	22 ABG29124	Novel human diagno
17	49	35.3	278	21 AAB56745	Human prostate can
18	49	35.3	400	22 ABB68410	Drosophila melanog
19	48	34.5	149	21 AAG27807	Arabidopsis thalia
20	48	34.5	188	21 AAG27806	Arabidopsis thalia
21	47.5	34.2	474	22 ABB62780	Drosophila melanog
22	47.5	34.2	474	22 ABB66823	Drosophila melanog
23	47	33.8	83	21 AAB34321	Human secreted pro
24	47	33.8	99	22 AAM39757	Human polypeptide
25	47	33.8	154	21 AAB25692	Human secreted pro
26	47	33.8	178	20 AAY41674	Human channel-rela
27	47	33.8	178	20 AAY13945	Human transmembran
28	47	33.8	178	22 AAU29294	Human PRO polypept
29	47	33.8	178	22 AAM38819	Human polypeptide
30	47	33.8	179	20 AAW88709	Secreted protein e
31	47	33.8	179	21 AAB25703	Human secreted pro
32	47	33.8	179	21 AAB25704	Human secreted pro
33	47	33.8	179	21 AAB25705	Human secreted pro
34	47	33.8	179	21 AAB25707	Human secreted pro
35	47	33.8	179	22 ABB50476	Human secreted pro
36	47	33.8	207	20 AAY40581	Partial amino acid
37	47	33.8	207	20 AAW85100	Thyroid hormone re
38	47	33.8	211	15 AAR53542	Thyroid hormone re
39	47	33.8	211	20 AAW92388	Human TR-interacti
40	47	33.8	219	22 ABB12045	Human channel-rela
41	47	33.8	219	22 AAM40605	Human polypeptide
42	47	33.8	220	20 AAW81640	Mouse elf-3 protei
43	47	33.8	273	21 AAB25706	Human secreted pro
44	47	33.8	635	21 AAB18176	Plasmodium falcipa
45	47	33.8	1120	20 AAW81642	Mouse elf protein.

ALIGNMENTS

RESULT 1
AAW56303
ID AAW56303 standard; Protein; 275 AA.
XX
AC AAW56303;
XX
DT 28-SEP-1998 (first entry)
XX
DE Babesia microti BMNI-20 antigen sequence.
XX
KW antigen; detection; diagnosis; vaccine; tick-borne disease;
KW differentiation; Lyme disease; ehrlichiosis.
XX
OS Babesia microti.
XX
PN EP834567-A2.
XX
PD 08-APR-1998.
XX
PF 01-OCT-1997; 97EP-0117067.
XX
PR 24-APR-1997; 97US-0845258.
PR 01-OCT-1996; 96US-0723142.
XX
PA (CORI-) CORIYA CORP.
XX
PI Houghton R, Lodes MJ, Reed SG, Sleath PR;
XX
DR WPI: 1998-195465/18.
XX
PT N-PSDB; AAV22749.
XX
PT Polypeptides comprising Babesia microti antigens and their
PT immunogenic fragments or epitopes - and related nucleic acid,
PT vectors, transformed cells and antibodies, useful for diagnosis of

PT infection and in protective vaccines
PS Claim 1; Page 101-102; 113pp; English.
XX
CC The sequence is that of a polypeptide comprising at least
CC one antigenic portion of a Babesia microti antigen. It can be used
CC to diagnose B. microti infection by detecting specific antibodies
CC in usual immunoassays. Infection can also be diagnosed using:
CC (a) primers or probes derived from the coding sequence, in
CC standard amplification or hybridisation tests, or (b) using
CC antibodies to detect the corresponding antigen. It is also
CC useful in vaccines to protect against infection, especially
CC when formulated with an adjuvant. The new diagnostic methods
CC allow rapid differentiation between B. microti infection and
CC other tick-borne diseases (Lyme disease and ehrlichiosis) that
CC have similar symptoms but require different treatments.
XX
SQ Sequence 275 AA;
Query Match 100.0%; Score 139; DB 19; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
DB 218 GKPNTNKSEKAERKSHDTQTTOEICE 243
RESULT 2
AAY24365
ID AAY24365 standard; Protein; 275 AA.
XX
AC AAY24365;
XX
DT 16-SEP-1999 (first entry)
XX
DE Babesia microti antigen BMNI-20 complementary open reading frame protein.
XX
KW Babesia microti; antigen; immunogen; diagnosis; infection; vaccine;
KW immunity; detection.
XX
OS Babesia microti.
XX
PN WO9929869-Al.
XX
PD 17-JUN-1999.
XX
PF 11-DEC-1998; 98WO-US26437.
XX
PR 11-DEC-1997; 97US-0990571.
XX
PA (CORI-) CORIXA CORP.
PA (MAYO-) MAYO FOUNDATION.
XX
PI Bruinsma E, Houghton R, Lodes MJ, Persing D, Reed SG;
PI Sleath PR;
XX
DR WPI; 1999-385612/32.
DR N-PSDB; AAX90017.
XX
PT New isolated Babesia microti polypeptides
PS Example 1; Page 109-110; 126pp; English.
XX
CC The present invention describes isolated polypeptides comprising
CC specific immunogenic portions of Babesia microti. AAX88983 to AAX88994
CC encode specifically claimed B. microti immunogenic proteins, and
CC AAY24327 to AAY24338 represent these proteins. B. microti polypeptides
CC and nucleic acids can be used for detecting B. microti infections. They
CC can also be used in vaccines for inducing protective immunity against B.
CC microti infections. The present sequence represents a B. microti antigen.
XX
SQ Sequence 275 AA;
Query Match 100.0%; Score 139; DB 19; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
DB 218 GKPNTNKSEKAERKSHDTQTTOEICE 243
RESULT 3
AAB30217
ID AAB30217 standard; Protein; 275 AA.
XX
AC AAB30217;
XX
DT 12-FEB-2001 (first entry)
XX
DE B. microti clone antigen SEQ ID NO: 53.
XX
KW Babesiosis; rodent parasite; tick-borne illness; antigen;
KW disease diagnosis; disease prevention.
XX
OS Babesia microti.
XX
PN WO200060090-Al.
XX
PD 12-OCT-2000.
XX
PF 05-APR-2000; 2000WO-US09136.
XX
PR 05-APR-1999; 99US-0286488.
PR 17-MAR-2000; 2000US-0528784.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Houghton RL, Sleath PR, McNeill PD;
XX
DR WPI; 2000-686939/67.
XX
PT New polypeptides containing an antigenic portion of Babesia microti
PT antigen and DNAs encoding the polypeptides, useful for diagnosing,
PT treating or preventing B. microti infection, or for inducing protective
PT immunity in a patient -
XX
PS Example 1; Page 98; 118pp; English.
XX
CC The present invention is related to the isolation of antigenic sequences
CC from the rodent parasite Babesia microti. This organism is transmitted to
CC humans by the same tick which transmits Lyme disease and ehrlichiosis.
CC The organism causes a malaria-like infection known as babesiosis. The
CC sequences identified by this invention can be used in the diagnosis,
CC prevention and treatment of babesiosis.
XX
SQ Sequence 275 AA;
Query Match 100.0%; Score 139; DB 21; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
DB 218 GKPNTNKSEKAERKSHDTQTTOEICE 243
RESULT 4
AAW56298
ID AAW56298 standard; Protein; 445 AA.
XX
AC AAW56298;
XX
DT 28-SEP-1998 (first entry)
XX

DE Babesia microti BMNI-17 complement antigen sequence.
 XX antigen; detection; diagnosis; vaccine; tick-borne disease;
 KW differentiation; Lyme disease; ehrlichiosis.
 XX

OS Babesia microti.
 XX

PN EP834567-A2.
 XX

PD 08-APR-1998.
 XX

XX 01-OCT-1997; 97EP-0117067.
 PF

XX 24-APR-1997; 97US-0845258.
 PR

PR 01-OCT-1996; 96US-0723142.
 XX

XX (CORI-) CORIXA CORP.
 PA

PI Houghton R, Lodes MJ, Reed SG, Sleath PR;
 XX

DR WPI; 1998-195465/18.
 XX

DR N-PSDB; AAV22753.
 XX

XX Polypeptides comprising Babesia microti antigens and their
 PT immunogenic fragments or epitopes - and related nucleic acid,
 PT vectors, transformed cells and antibodies, useful for diagnosis of
 PT infection and in protective vaccines
 PT

XX Claim 1; Page 77-79; 113pp; English.
 PS

XX The sequence is that of a polypeptide comprising at least
 CC one antigenic portion of a Babesia microti antigen. It can be used
 CC to diagnose B. microti infection by detecting specific antibodies
 CC in usual immunoassays. Infection can also be diagnosed using:
 CC (a) primers or probes derived from the coding sequence, in
 CC standard amplification or hybridisation tests, or (b) using
 CC antibodies to detect the corresponding antigen. It is also
 CC useful in vaccines to protect against infection, especially
 CC when formulated with an adjuvant. The new diagnostic methods
 CC allow rapid differentiation between B. microti infection and
 CC other tick-borne diseases (Lyme disease and ehrlichiosis) that
 CC have similar symptoms but require different treatments.
 CC

XX Sequence 445 AA;
 SQ

Query Match 100.0%; Score 139; DB 19; Length 445;
 Best Local Similarity 100.0%; Pred. No. 5.9e-13;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKARSHDTQTTOEICE 26
 |||||

Db 182 GKPNTNKSEKARSHDTQTTOEICE 207
 |||||

RESULT 5

AAV24358

ID AAV24358 standard; Protein; 445 AA.
 XX

AC AAV24358;
 XX

DT 16-SEP-1999 (first entry)
 XX

XX Babesia microti antigen BMNI-17 complementary open reading frame protein.
 DE

XX Babesia microti; antigen; immunogen; diagnosis; infection; vaccine;
 KW immunity; detection.
 XX

OS Babesia microti.
 XX

PN W09929869-A1.
 XX

XX 17-JUN-1999.
 PS

XX

PF 11-DEC-1998; 98WO-US26437.
 XX

PR 11-DEC-1997; 97US-0990571.
 XX

XX (CORI-) CORIXA CORP.
 PA

PA (MAYO-) MAYO FOUNDATION.
 XX

PI Bruinsma E, Houghton R, Lodes MJ, Persing D, Reed SG;
 Sleath PR;
 XX

XX WPI; 1999-385612/32.
 DR

DR N-PSDB; AAX90012.
 XX

XX New isolated Babesia microti polypeptides
 PT

XX Example 1; Page 91-92; 126pp; English.
 PS

XX The present invention describes isolated polypeptides comprising
 CC specific immunogenic portions of Babesia microti. AAX88983 to AAX8994
 CC encode specifically claimed B. microti immunogenic proteins, and
 CC AAY24327 to AAY24338 represent these proteins. B. microti polypeptides
 CC and nucleic acids can be used for detecting B. microti infections. They
 CC can also be used in vaccines for inducing protective immunity against B.
 CC microti infections. The present sequence encodes a B. microti antigen.
 XX

SQ Sequence 445 AA;
 SQ

Query Match 100.0%; Score 139; DB 20; Length 445;
 Best Local Similarity 100.0%; Pred. No. 5.9e-13;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKARSHDTQTTOEICE 26
 |||||

Db 182 GKPNTNKSEKARSHDTQTTOEICE 207
 |||||

RESULT 6

AAB30207

ID AAB30207 standard; Protein; 445 AA.
 XX

AC AAB30207;
 XX

DT 12-FEB-2001 (first entry)
 XX

XX B. microti BMNI-17 antigen reverse complement SEQ ID NO: 38.
 DE

XX Babesiosis; rodent parasite; tick-borne illness; antigen;
 KW disease diagnosis; disease prevention.
 XX

OS Babesia microti.
 XX

PN W0200060090-A1.
 XX

PD 12-OCT-2000.
 XX

XX 05-APR-2000; 2000WO-US09136.
 PF

XX 05-APR-1999; 99US-0286488.
 PR

PR 17-MAR-2000; 2000US-0528784.
 XX

XX (CORI-) CORIXA CORP.
 PA

XX Reed SG, Lodes MJ, Houghton RL, Sleath PR, McNeill PD;
 PI

XX WPI; 2000-686939/67.
 DR

XX New polypeptides containing an antigenic portion of Babesia microti
 PT antigen and DNAs encoding the polypeptides, useful for diagnosing,
 PT treating or preventing B. microti infection, or for inducing protective
 PT immunity in a patient -
 XX

XX Example 1; Page 86-87; 118pp; English.
 PS

XX

CC The present invention is related to the isolation of antigenic sequences
 CC from the rodent parasite Babesia microti. This organism is transmitted to
 CC humans by the same tick which transmits Lyme disease and ehrlichiosis.
 CC The organism causes a malaria-like infection known as babesiosis. The
 CC sequences identified by this invention can be used in the diagnosis,
 CC prevention and treatment of babesiosis.
 XX
 XX

SQ Sequence 445 AA;

Query Match 100.0%; Score 139; DB 21; Length 445;
 Best Local Similarity 100.0%; Pred. No. 5.9e-13;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
 DB 182 GKPNTNKSEKAERKSHDTQTTOEICE 207

RESULT 7
 AAB30230
 ID AAB30230 standard; Protein; 666 AA.

XX
 AC AAB30230;

XX
 DT 12-FEB-2001 (first entry)

XX B. microti MN-10/BMNI-17 fusion protein SEQ ID NO: 85.

XX Babesiosis; rodent parasite; tick-borne illness; antigen;
 KW disease diagnosis; disease prevention.

XX Babesia sp.
 OS Synthetic.

XX WO200060090-A1.

XX 12-OCT-2000.

XX 05-APR-2000; 2000WO-US09136.

XX 05-APR-1999; 99US-0286488.

XX 17-MAR-2000; 2000US-0528784.

XX (CORI-) CORIXA CORP.

XX Reed SG, Lodes MJ, Houghton RL, Sleath PR, McNeill PD;

XX WPI; 2000-686939/67.

XX New polypeptides containing an antigenic portion of Babesia microti
 PT antigen and DNAs encoding the polypeptides, useful for diagnosing,
 PT treating or preventing B. microti infection, or for inducing protective
 PT immunity in a patient -

XX Example 7; Page 108-111; 118pp; English.

XX The present invention is related to the isolation of antigenic sequences
 CC from the rodent parasite Babesia microti. This organism is transmitted to
 CC humans by the same tick which transmits Lyme disease and ehrlichiosis.
 CC The organism causes a malaria-like infection known as babesiosis. The
 CC sequences identified by this invention can be used in the diagnosis,
 CC prevention and treatment of babesiosis.

XX Sequence 666 AA;

Query Match 100.0%; Score 139; DB 21; Length 666;
 Best Local Similarity 100.0%; Pred. No. 9.4e-13;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
 DB 403 GKPNTNKSEKAERKSHDTQTTOEICE 428

7

RESULT 8
 AAB30231
 ID AAB30231 standard; Protein; 1132 AA.

XX
 AC AAB30231;

XX 12-FEB-2001 (first entry)

XX B. microti MN-10/BMNI-17/BMNI-15 fusion protein SEQ ID NO: 87.
 XX Babesiosis; rodent parasite; tick-borne illness; antigen;
 KW disease diagnosis; disease prevention.

XX Babesia sp.
 OS Synthetic.

XX WO200060090-A1.

XX 12-OCT-2000.

XX 05-APR-2000; 2000WO-US09136.

XX 05-APR-1999; 99US-0286488.

XX 17-MAR-2000; 2000US-0528784.

XX (CORI-) CORIXA CORP.

XX Reed SG, Lodes MJ, Houghton RL, Sleath PR, McNeill PD;

XX WPI; 2000-686939/67.

XX New polypeptides containing an antigenic portion of Babesia microti
 PT antigen and DNAs encoding the polypeptides, useful for diagnosing,
 PT treating or preventing B. microti infection, or for inducing protective
 PT immunity in a patient -

XX Example 7; Page 112-116; 118pp; English.

XX The present invention is related to the isolation of antigenic sequences
 CC from the rodent parasite Babesia microti. This organism is transmitted to
 CC humans by the same tick which transmits Lyme disease and ehrlichiosis.
 CC The organism causes a malaria-like infection known as babesiosis. The
 CC sequences identified by this invention can be used in the diagnosis,
 CC prevention and treatment of babesiosis.

SQ Sequence 1132 AA;

Query Match 100.0%; Score 139; DB 21; Length 1132;
 Best Local Similarity 100.0%; Pred. No. 1.7e-12;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
 DB 869 GKPNTNKSEKAERKSHDTQTTOEICE 894

RESULT 9
 AAY24359

ID AAY24359 standard; peptide; 32 AA.

XX
 AC AAY24359;

XX 16-SEP-1999 (first entry)

XX Babesia microti antigen BMNI-17 degenerate repeat sequence.

XX Babesia microti; antigen; immunogen; diagnosis; infection; vaccine;
 KW immunity; detection.

XX Babesia microti.

XX Key Location/Qualifiers

FT Misc-difference 3 /label= Gly, Asp
FT Misc-difference 5 /label= Pro, Ile
FT Misc-difference 70 /label= Lys, Thr
FT Misc-difference 11 /label= Glu, Gly
FT Misc-difference 12 /label= Lys, Asn
FT Misc-difference 14 /label= Glu, Gly
FT Misc-difference 15 /label= Ile, Arg
FT Misc-difference 18 /label= His, Tyr
FT Misc-difference 23 /label= Thr, Pro
FT Misc-difference 26 /label= Ile, Thr
FT Misc-difference 27 /label= Cys, Ser
FT Misc-difference 28 /label= Asp, Glu
FT Misc-difference 29 /label= Glu, Ala
FT Misc-difference 30 /label= Cys, His
XX WO9929869-A1.
PN
XX
XX 17-JUN-1999.
XX
XX 11-DEC-1998; 98WO-US26437.
XX
XX 11-DEC-1997; 97US-0990571.
PR
XX (CORI-) CORIXA CORP.
PA (MAYO-) MAYO FOUNDATION.
PA
XX Bruinsma E, Houghton R, Lodes MJ, Persing D, Reed SG;
PI Sleath PR;
XX
XX WPI; 1999-385612/32.
XX
XX New Isolated Babesia microti polypeptides
XX
XX Example 1; Page 94; 126pp; English.
XX
CC The present invention describes isolated polypeptides comprising
CC specific immunogenic portions of Babesia microti. AAX8893 to AAX8894
CC encode specifically claimed B. microti immunogenic proteins, and
CC AAY24327 to AAY24338 represent these proteins. B. microti polypeptides
CC and nucleic acids can be used for detecting B. microti infections. They
CC can also be used in vaccines for inducing protective immunity against B.
CC microti infections. The present sequence represents a B. microti antigen
CC BMNI-17 degenerate repeat sequence.
XX
XX Sequence 32 AA;
SQ
Query Match 45.3%; Score 63; DB 20; Length 32;
Best Local Similarity 63.6%; Pred. No. 0.013;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 2 KPNTNKSEKAEKSHDTQTQOE 23
DB 4 KXNKNKXXXXXKSDTQTQOE 25
RESULT 10
AAB30208
ID AAB30208 standard; Peptide; 32 AA.
XX

AC AAB30208;
XX
XX 12-FEB-2001 (first entry)
XX
DE B. microti BMNI-17 antigen reverse complement repeat SEQ ID NO: 39.
XX
XX Babesiosis; rodent parasite; tick-borne illness; antigen;
KW disease diagnosis; disease prevention.
KW
XX Babesia microti.
OS
XX WO200060090-A1.
PN
XX 12-OCT-2000.
XX
XX 05-APR-2000; 2000WO-US09136.
XX
XX 05-APR-1999; 99US-0286488.
PR
XX 17-MAR-2000; 2000US-0528784.
XX
XX (CORI-) CORIXA CORP.
PA
XX Reed SG, Lodes MJ, Houghton RL, Sleath PR, McNeill PD;
PI
XX WPI; 2000-686939/67.
DR
XX
XX New polypeptides containing an antigenic portion of Babesia microti
PT antigen and DNAs encoding the polypeptides, useful for diagnosing,
PT treating or preventing B. microti infection, or for inducing protective
PT immunity in a patient -
XX
XX Claim 6; Page 88; 118pp; English.
PS
XX The present invention is related to the isolation of antigenic sequences
CC from the rodent parasite Babesia microti. This organism is transmitted to
CC humans by the same tick which transmits Lyme disease and ehrlichiosis.
CC The organism causes a malaria-like infection known as babesiosis. The
CC sequences identified by this invention can be used in the diagnosis,
CC prevention and treatment of babesiosis.
XX
XX Sequence 32 AA;
SQ
Query Match 45.3%; Score 63; DB 21; Length 32;
Best Local Similarity 63.6%; Pred. No. 0.013;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 2 KPNTNKSEKAEKSHDTQTQOE 23
DB 4 KXNKNKXXXXXKSDTQTQOE 25
RESULT 11
AAG75092
ID AAG75092 standard; Protein; 600 AA.
XX
XX AAG75092;
AC
XX 03-SEP-2001 (first entry)
DT
XX Human colon cancer antigen protein SEQ ID NO:5856.
DE
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 21.
KW
XX Homo sapiens.
OS
XX WO200122920-A2.
PN
XX 05-APR-2001.
PD
XX 28-SEP-2000; 2000WO-US26524.
PF
XX 29-SEP-1999; 99US-0157137.
PR

```
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI; 2001-235357/24.
DR N-PSDB; AAH34497.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PT
XX Claim 11; Page 7370-7373; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 600 AA;
SQ
Query Match 38.8%; Score 54; DB 22; Length 600;
Best Local Similarity 46.2%; Pred. No. 9;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
QY 1 GKPNTNKEKAERKSHDQTQICE 26
DB 351 GNPNGKSEKKEKTPLRDESTQEHSE 376
IIIIII : : : : :
IIIIII : : : : :
RESULT 12
AAY68823
ID AAY68823 standard; protein; 1048 AA.
XX
XX AAY68823;
AC
XX 16-MAY-2000 (first entry)
DT
XX Amino acid sequence of a Son of sevenless (Sos) related protein.
DE
XX Ras; Son of sevenless; Sos; crystal; Ras-Sos complex; cancer;
KW protein coordinate data.
XX
XX Saccharomyces cerevisiae.
OS
XX WO200005258-A1.
PN
XX 03-FEB-2000.
PD
XX 20-JUL-1999; 99WO-US16348.
PF
XX 21-JUL-1998; 98US-0119794.
PR
XX (UYRQ ) UNIV ROCKEFELLER.
PA
XX Boriack-Sjodin A, Margarit SM, Bar-Sagi D, Cole P, Kuriyan J;
PI WPI; 2000-182647/16.
XX
XX
```

```
XX Novel crystals comprising a Ras-Son of sevenless complex, useful for
PT screening drugs useful in cancer treatment -
PT
XX Disclosure; Page 181-185; 224pp; English.
XX
XX The specification describes a crystal complex comprising at least a
CC Sos contacting region of a Ras protein and at least a Son of sevenless
CC (Sos) protein catalytic region fragment, that effectively diffracts
CC X-rays. Ras and Sos form a tight complex, Sos does not impede the
CC binding sites for the nucleotide base and the ribose of GTP or GDP
CC and thus the Ras-Sos complex maintains a structure that permits
CC nucleotide release and rebinding. The crystals are used for the
CC determination of the atomic coordinates of the complex to a resolution
CC of more than 5.0 Angstrom. The crystals, or a dataset comprising the
CC three-dimensional coordinates obtained from the crystals, is useful
CC for identifying an agent that stabilizes the Ras-Sos complex. The
CC crystals are also useful for identifying agents that inhibit the
CC formation of Ras-Sos complex. Ras and Sos fragments are useful for
CC growing a crystal of a protein-ligand complex. Agents that stabilize
CC or inhibit the formation of Ras-Sos complex are useful in the
CC treatment of cancer. The present sequence represents a yeast
CC Sos-related protein.
XX
XX Sequence 1048 AA;
SQ
Query Match 38.8%; Score 54; DB 21; Length 1048;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 3 PNTNKEKAERKSHDQTQTTQ 22
DB 1002 PNSNSNNKSEKSRDDQTDE 1021
IIIIII : : : : :
IIIIII : : : : :
RESULT 13
AAY35210
ID AAY35210 standard; Protein; 212 AA.
XX
XX AAY35210;
AC
XX 13-SEP-1999 (first entry)
DT
XX Chlamydia pneumoniae protein not found in C. trachomatis.
DE
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
XX Chlamydia pneumoniae.
OS
XX WO9927105-A2.
PN
XX 03-JUN-1999.
PD
XX 20-NOV-1998; 98WO-IB01890.
PF
XX 04-NOV-1998; 98US-0107078.
PR
XX 21-NOV-1997; 97FR-0014673.
XX
XX (GEST ) GENSET.
PA
XX Griffiths R;
XX
XX WPI; 1999-357842/30.
DR
XX Genome sequence of Chlamydia pneumoniae
PT
XX Page 1060; Disclosure; 1912pp; English.
PS
XX AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC
```

CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae
XX
SQ Sequence 212 AA;

Query Match 36.7%; Score 51; DB 20; Length 212;
Best Local Similarity 38.1%; Pred. No. 8;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 6 NKSEKAERKSHDTQTQICE 26
|:|:|::|:|:|::|:|:|
Db 19 NQORDIQRDREASQACE 39

RESULT 14
AAY34934
ID AAY34934 standard; Protein; 584 AA.

XX AC AAY34934;

XX 13-SEP-1999 (first entry)

XX Chlamydia pneumoniae transmembrane protein sequence.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

XX WO9927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-IB01890.

XX 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

XX (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

XX Page 860-861; Disclosure; 1912pp; English.

XX AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 584 AA;

Query Match 36.0%; Score 50; DB 20; Length 584;
Best Local Similarity 30.8%; Pred. No. 36;
Matches 8; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GKPNTNKSEKAERKSHDTQTQICE 26
|:|:|::|:|:|::|:|:|
Db 292 GIPNSNTLERAEKAEKQESREOLSE 317

RESULT 15

AAY56856

ID AAY56856 standard; Protein; 523 AA.

XX AC AAY56856;

XX 10-APR-2000 (first entry)

XX Morbillivirus antigenic protein.

XX Antigenic; nucleocapsid protein; canine distemper virus; CDV;

XX Morbillivirus; Paramyxovirus.

XX Morbillivirus sp.

XX JPI1346768-A.

XX 21-DEC-1999.

XX 03-JUN-1998; 98JP-0155072.

XX 03-JUN-1998; 98JP-0155072.

XX (MORG) MORINAGA MILK IND CO LTD.

XX WPI; 2000-109685/10.

DR N-PSDB; AAZ46837.

XX A protein having antigenicity of canine distemper virus nucleocapsid

XX protein - useful in the diagnosis of canine distemper virus infection

XX Disclosure; Page 11-13; 15pp; Japanese.

XX The invention provides a protein which has a molecular weight of 58 kD
CC and has antigenicity of nucleocapsid protein (NP) of canine distemper
CC virus (CDV) belonging to Morbillivirus genus of Paramyxovirus family.
CC The protein can be used to prepare a reagent for the determination of
CC anti-CDV NP antibody by immobilizing the above protein on a carrier or a
CC membrane as the active component. The reagent can be used for diagnosis
CC of CDV infection, judgement of effect of vaccine and judgement of
CC inoculation period. The reagent can determine anti-CDV NP antibody
CC contained in canine serum easily. The present sequence represents
CC the Morbillivirus antigenic protein.

XX Sequence 523 AA;

Query Match 35.6%; Score 49.5; DB 21; Length 523;
Best Local Similarity 47.6%; Pred. No. 38;
Matches 10; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 3 PNTNKSEKAERKSHDTQTQICE 23

|:|:|::|:|:|::|:|:|

Db 461 PDVNSERSEPR-HDTQIIQD 480

Search completed: October 11, 2002, 20:37:37
Job time : 64 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

protein - protein search, using sw model

Run on: October 11, 2002, 20:36:38 ; Search time 33 seconds
(without alignments)
19.244 Million cell updates/sec

Title: US-09-794-764-195

Perfect score: 139

Sequence: 1 GRPNTNKEAERKSHDTQTTOEICE 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 145 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	100.0	275	4	Sequence 53, Appl
2	139	100.0	275	4	Sequence 53, Appl
3	139	100.0	445	4	Sequence 38, Appl
4	139	100.0	445	4	Sequence 38, Appl
5	139	100.0	445	4	Sequence 38, Appl
6	63	45.3	32	4	Sequence 39, Appl
7	63	45.3	32	4	Sequence 39, Appl
8	63	45.3	32	4	Sequence 39, Appl
9	54	38.8	1048	3	Sequence 39, Appl
10	47	33.8	220	2	Sequence 5, Appl
11	47	33.8	220	2	Sequence 5, Appl
12	46	33.1	259	4	Sequence 4, Appl
13	46	33.1	259	4	Sequence 4, Appl
14	44	31.7	343	1	Sequence 13, Appl
15	44	31.7	343	1	Sequence 10, Appl
16	43	30.9	123	5	Sequence 10, Appl
17	43	30.9	123	5	Sequence 17, Appl
18	43	30.9	123	5	Sequence 19, Appl
19	43	30.9	731	4	Sequence 8, Appl
20	43	30.9	752	1	Sequence 11, Appl
21	42	30.2	461	2	Sequence 2, Appl
22	42	30.2	461	2	Sequence 26, Appl
23	42	30.2	461	3	Sequence 4, Appl
24	42	30.2	461	5	Sequence 26, Appl
25	42	30.2	546	4	Sequence 27, Appl
26	42	30.2	588	2	Sequence 12, Appl
27	42	30.2	588	3	Sequence 12, Appl

US-08-845-258-53
; Sequence 53, Application US/08845258
; Patent No. 6183976
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,258
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Babesia Microti
; US-08-845-258-53

ALIGNMENTS

RESULT 1
US-08-845-258-53
; Sequence 53, Application US/08845258
; Patent No. 6183976
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,258
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Babesia Microti
; US-08-845-258-53

Query Match 100.0%; Score 139; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
|||||
Db 218 GKPNTNKSEKAERKSHDTQTTOEICE 243

RESULT 2

US-08-990-571-53
; Sequence 53, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B. M
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,571
; FILING DATE: 11-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Babesia Microti
US-08-990-571-53

Query Match 100.0%; Score 139; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
|||||
Db 218 GKPNTNKSEKAERKSHDTQTTOEICE 243

RESULT 3

US-08-845-258-38
; Sequence 38, Application US/08845258
; Patent No. 6183976
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; AND TREATMENT OF B. MICROTI INFECTION
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington

; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,258
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 445 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-845-258-38

Query Match 100.0%; Score 139; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 2.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKPNTNKSEKAERKSHDTQTTOEICE 26
|||||
Db 182 GKPNTNKSEKAERKSHDTQTTOEICE 207

RESULT 4

US-08-990-571-38
; Sequence 38, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B.
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,571
; FILING DATE: 11-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 445 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-990-571-38

; OTHER INFORMATION: /note= "Residue can be either Thr
; OTHER INFORMATION: or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 26
; OTHER INFORMATION: /note= "Residue can be either Ile
; OTHER INFORMATION: or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 27
; OTHER INFORMATION: /note= "Residue can be either Cys
; OTHER INFORMATION: or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 28
; OTHER INFORMATION: /note= "Residue can be either Asp
; OTHER INFORMATION: or Glu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 29
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Ala"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 30
; OTHER INFORMATION: /note= "Residue can be either Cys
; OTHER INFORMATION: or His"
US-08-845-258-39

Query Match 45.3%; Score 63; DB 4; Length 32;
Best Local Similarity 63.6%; Pred. No. 0.004;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 2 KPTNKKSEKAERKSHDTQTQCE 23
| | | | | | | | | | | | | | | | | | | |
Db 4 KXNKNKXXXXXSDTQTQXE 25

RESULT 7
US-08-990-571-39
; Sequence 39, Application us/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B. M
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,571
; FILING DATE: 11-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid

; STRANDEDNESS:
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /note= "Residue can be either Gly
; OTHER INFORMATION: or Asp"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 5
; OTHER INFORMATION: /note= "Residue can be either Pro
; OTHER INFORMATION: or Ile"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 7
; OTHER INFORMATION: /note= "Residue can be either Lys
; OTHER INFORMATION: or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 12
; OTHER INFORMATION: /note= "Residue can be either Lys
; OTHER INFORMATION: or Asn"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 14
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 15
; OTHER INFORMATION: /note= "Residue can be either Ile
; OTHER INFORMATION: or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18
; OTHER INFORMATION: /note= "Residue can be either His
; OTHER INFORMATION: or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 23
; OTHER INFORMATION: /note= "Residue can be either Thr
; OTHER INFORMATION: or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 26
; OTHER INFORMATION: /note= "Residue can be either Ile
; OTHER INFORMATION: or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 27
; OTHER INFORMATION: /note= "Residue can be either Cys
; OTHER INFORMATION: or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 28
; OTHER INFORMATION: /note= "Residue can be either Asp
; OTHER INFORMATION: or Glu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 29
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Ala"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 30
; OTHER INFORMATION: /note= "Residue can be either Cys
; OTHER INFORMATION: or His"
US-08-990-571-39

Query Match 45.3%; Score 63; DB 4; Length 32;
Best Local Similarity 63.6%; Pred. No. 0.004;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 KPNTNKSEAKRSHDTQTQOE 23
| | | | | | | | | | | | | | | | | | | |
Db 4 KXNKNKXXXXXKSDTQTQOE 25

RESULT 8

US-08-723-142A-39
; Sequence 39, Application US/08723142A
; Patent No. 6306396
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,142A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /note= "Residue can be either Gly
; OTHER INFORMATION: or Asp"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 5
; OTHER INFORMATION: /note= "Residue can be either Pro
; OTHER INFORMATION: or Ile"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 7
; OTHER INFORMATION: /note= "Residue can be either Lys
; OTHER INFORMATION: or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Gly"
; FEATURE:
; NAME/KEY: Modified-site

; LOCATION: 12
; OTHER INFORMATION: /note= "Residue can be either Lys
; OTHER INFORMATION: or Asn"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 14
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 15
; OTHER INFORMATION: /note= "Residue can be either Ile
; OTHER INFORMATION: or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18
; OTHER INFORMATION: /note= "Residue can be either His
; OTHER INFORMATION: or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 23
; OTHER INFORMATION: /note= "Residue can be either Thr
; OTHER INFORMATION: or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 26
; OTHER INFORMATION: /note= "Residue can be either Ile
; OTHER INFORMATION: or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 27
; OTHER INFORMATION: /note= "Residue can be either Cys
; OTHER INFORMATION: or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 28
; OTHER INFORMATION: /note= "Residue can be either Asp
; OTHER INFORMATION: or Glu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 29
; OTHER INFORMATION: /note= "Residue can be either Glu
; OTHER INFORMATION: or Ala"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 30
; OTHER INFORMATION: /note= "Residue can be either Cys
; OTHER INFORMATION: or His"
; US-08-723-142A-39

Query Match 45.3%; Score 63; DB 4; Length 32;
Best Local Similarity 63.6%; Pred. No. 0.004;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 KPNTNKSEAKRSHDTQTQOE 23
| | | | | | | | | | | | | | | | | | | |
Db 4 KXNKNKXXXXXKSDTQTQOE 25

RESULT 9

US-09-356-952-5
; Sequence 5, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borlack-Sjodin, Ann
; APPLICANT: Margarit, S. M.
; APPLICANT: Bor-Sogli, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952

ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/185.160
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/065,459
FILING DATE: NOVEMBER 12, 1997
ATTORNEY/AGENT INFORMATION:
NAME: MAJARIAN, WILLIAM R.
REGISTRATION NUMBER: 41,173
REFERENCE/DOCKET NUMBER: BB-1096
TELEPHONE: 302-992-4926
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 734 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-185-160-13

Query Match 33.1%; Score 46; DB 4; Length 734;
Best Local Similarity 42.1%; Pred. No. 56;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 KPNTNKAERKSHDTOT 20
||| | | | | | | | | |
DB 687 KPETHKAGKSNQHTGT 705

RESULT 14
US-08-348-792-10
; Sequence 10, Application US/08348792
; Patent No. 5576423
; GENERAL INFORMATION:
; APPLICANT: Aversa, Gregorio
; APPLICANT: Chang, Chia-Chun J.
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: de Vries, Jan E.
; TITLE OF INVENTION: PURIFIED GENES ENCODING MAMMALIAN CELL
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,792
FILING DATE: 02-DEC-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0436
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-348-792-10
Query Match 31.7%; Score 44; DB 1; Length 343;
Best Local Similarity 47.4%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 7 KSEKAERKSHDTQTQEQIC 25
||| | | | | | | | | |
DB 293 KSGPQEKHLHALTDQDPC 311
RESULT 15
US-08-462-738-10
; Sequence 10, Application US/08462738
; Patent No. 5977303
; GENERAL INFORMATION:
; APPLICANT: Aversa, Gregorio
; APPLICANT: Chang, Chia-Chun J.
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: de Vries, Jan E.
; TITLE OF INVENTION: PURIFIED GENES ENCODING MAMMALIAN CELL
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,738
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,792
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0436GB
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-738-10

Query Match 31.7%; Score 44; DB 2; Length 343;
Best Local Similarity 47.4%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 7 KSEKAERKSHDTQTQEQIC 25
||| | | | | | | | | |
DB 293 KSGPQEKHLHALTDQDPC 311

Job time : 34 secs

THIS PAGE BLANK (USPTO)